AD_____

GRANT NUMBER DAMD17-96-1-6234

TITLE: Role of Ser-Arg Proteins in the Regulation of RNA Processing

PRINCIPAL INVESTIGATOR: Benjamin J. Blencowe, Ph.D.

CONTRACTING ORGANIZATION: Massachusetts Institute of Technology
Cambridge, Masschusetts 02139

REPORT DATE: September 1997

DTIC QUALITY INSPECTED 2

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

19980317 151

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE September 1997	3. REPORT TYPE AND DATES COVERED Annual (1 Sep 96 - 31 Aug 97)

4. TITLE AND SUBTITLE
Role of Ser-Arg Proteins in the Regulation of RNA Processing

5. FUNDING NUMBERS
DAMD17-96-1-6234

6. AUTHOR(S)
Benjamin J. Blencowe, Ph.D

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200)*

Deregulation of alternative splicing has been linked to malignant transformation in breast cancers. Therefore, to fully understand breast cancer, it will be important to identify and characterize factors that regulate the splicing process. Nuclear matrix proteins (NMp's) containing serine/arginine (SR)-rich domains that associate with splicing complexes were isolated and characterized. The NMp's B1C8 (160kDa) and B4A11 (300kDa) are novel SR phosphoproteins of 820 and 2245 amino acids, respectively. Unlike previously defined SR family proteins (SRp's), which function as constitutive and regulatory splicing factors, B1C8 and B4A11 lack RNA Recognition Motifs (RRMs). B1C8 and B4A11 form a complex (B1C8/B4A11) that associates with SR family proteins SRp40 and SRp75. B1C8/B4A11 is required for the splicing of a subset of pre-mRNAs in vitro. B1C8/B4A11 binds to pre-mRNA and promotes splicing in combination with SR family proteins. The association of B1C8/B4A11 with pre-mRNA also requires binding of U1 snRNP to the 5' splice site and is stabilized by binding of U2 snRNP to the branch site. Thus, B1C8/B4A11 appears to form critical interactions that span intron sequences. These results suggest that B1C8/B4A11 functions as a splicing "coactivator", by augmenting the activity of factors bound to pre-mRNA. These results are providing new insights into mechanisms by which pre-mRNA splicing may be regulated.

14. SUBJECT TERMS Breast Cancer		15. NUMBER OF PAGES 36
		16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

_____ Where copyrighted material is quoted, permission has been obtained to use such material.

_____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

_____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

_BB_ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

_____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

_BB_ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

_BB_ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

_____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Ben Blenowe ___ 10/15/97
PI - Signature _____ Date

CONTENTS   **PAGE**

Introduction

Deregulation of alternative splicing as been linked to malignant transformation and the formation of metastases in breast cancers. For example, specific alternative spliced forms of the cell surface adhesion glycoprotein CD44 have been correlated with invasive tumor formation. Moreover, it has been demonstrated that expression of specific alternative spliced forms of CD44 mRNAs in non-metastatic cell lines result in transition to full metastatic potential. We are interested in understanding the mechanisms underlying the regulation of alternative splicing with the long term goal of identifying and targeting trans-acting splicing factors that are involved in metastatic transitions in breast and other types of cancer.

Precursor mRNA (pre-mRNA) splicing occurs within the spliceosome, a ~60S complex composed of four small nuclear ribonucleoprotein particles (U1, U2, U4/U6, U5 snRNPs) and multiple non-snRNP factors (reviewed by Sharp, 1994; Krämer, 1996). Although a wealth of biochemical and genetic data has provided detailed information on the role of specific spliceosomal components, little is known about the recognition of exon-intron segments in vivo, or how alternative splicing is regulated.

Increasing evidence suggests that a family of highly conserved non-snRNP splicing factors containing sequences rich in alternating serine and arginine residues (SR proteins) may function in the majority, if not all, constitutive and regulated pre-mRNA splicing events in multicellular organisms. The SR-family of proteins so far comprises a group of eight factors (SRp20, SRp30a/ASF/SF2, SRp30b/SC35, SRp30c, 9G8, SRp40, SRp55 and SRp75 (reviewed by Fu, 1995; Manley and Tacke, 1996; Reed, 1996; Valcárcel and Green, 1996). Members of this family share a similar structural organization, consisting of one or two N-terminal RNA Recognition Motifs (RRMs) and an extensively phosphorylated C-terminal domain that is rich in SR dipeptides (RS domain). The RRMs of SR proteins bind RNA with some sequence-specificity, whereas the RS domains primarily function in protein-protein interactions with other SR proteins.

In recent studies supported by the Breast Cancer Research Program, we have isolated and characterized two nuclear matrix proteins involved in pre-mRNA splicing, B1C8 (160kDa) and B4A11 (300kDa) (Blencowe et al., 1994, 1995). B1C8 and B4A11 are novel proteins that contain multiple SR repeats but, unlike members of the SR-family of splicing factors, lack RRMs (Blencowe et al., submitted; unpublished observations). B1C8 and B4A11 form a complex (B1C8/B4A11) that associates with a subset of the SR-family proteins. B1C8/B4A11 binds to pre-mRNA and promotes splicing activity through co-operative interactions with SR proteins. Interestingly however, B1C8/B4A11 is only required for splicing of specific pre-mRNAs. These and additional studies summarized in this year's annual report indicate that B1C8/B4A11 functions as a splicing "coactivator". It appears to promote splicing by bridging between sequence-specific splicing factors and other pre-mRNA-bound components required for spliceosome formation (Blencowe et al., submitted). Additional highlight's of progress made during the past year include the isolation and complete sequencing of cDNAs encoding the 2225 amino acid B4A11 matrix protein and the identification of specific interactions between B1C8/B4A11 and splicing snRNPs bound to pre-mRNA.

The studies summarized in this report have, for the first time, provided information on the roles of integral nuclear matrix proteins in splicing. They have also provided a working model for how specific sets of exons may be recognized and paired during splicing. It is anticipated that these insights will pave the way for future investigations of novel mechanisms underlying the regulation of splice site selection, in both normal and malignant

4

cells. They also provide an entrée into future investigations of the role of nuclear matrix proteins in the organization of pre-mRNA processing in vivo.

BODY

In last year's report, the isolation and preliminary characterization of the B1C8/B4A11 complex of nuclear matrix-SR proteins was described. These studies demonstrated that B1C8/B4A11 promotes splicing activity through cooperative interactions with SR family proteins. Research during the past year, as part of Year 1 of the Career Development Award (CDA), has extended these studies. For the sake of clarity, the new results will be described in the context of previous work. A set of data figures follows the body of the report. A discussion of how the results of the past year address the Statement of Work (SoW) for year 1 of the CDA follows the report body.

Isolation of cDNAs encoding the B1C8 and B4A11 matrix proteins

Previously, B1C8-160kDa and B4A11-300kDa were partially purified by a three-step procedure using conventional chromatography and magnesium precipitation, as previously described (Blencowe et al., 1995). The proteins were recovered from a SDS polyacrylamide gel, digested with lys-C, and resulting peptides were microsequenced. Searches of data bases with the microsequences did not identify known proteins although microsequences were contained within a partial human cDNA sequences in the data base of expressed sequence tags (ESTs). Fragments of ESTs cDNAs corresponding to the peptide sequences were used to screen a λ cDNA library derived from human U937 cells and positive plaques were isolated. Sequence analysis of the longest cDNA (3.7 kb) isolated from one of the plaques, and of several overlapping EST clones identified by data base searches, revealed a single B1C8 ORF of 820 amino acid residues. Similarly, during the past year, assembly of overlapping cDNAs to generate a ~10kb cDNA sequence revealed an uninterrupted ORF of 2225 amino acid resides corresponding to B4A11. All of the B1C8-160kDa and B4A11-300kDa peptide microsequences were contained within these respective ORFs. A schematic diagram of the B1C8 and B4A11 ORFs is shown in Figure 1.

Northern analysis of polyadenylated RNA isolated from different human tissues with a cDNA probe corresponding to an N-terminal B1C8 ORF sequences detected a major ~3.8 kb transcript and a second transcript of ~6.0kb (Figure 1B). The ~3.8 kb transcript is ubiquitous although variably expressed whereas the 6.0 kb species appears to be expressed only in a subset of human tissues. Interestingly, a similar pattern of variable expression was observed for a ~10kb transcript detected with a cDNA probe to an N-terminal B4A11 ORF sequence (data not shown). Each transcript was approximately the same length as the ORF plus the 5' and 3' UTRs in the longest cDNA inserts. The first in-frame methionine residues of both ORFs were preceded by in frame stop codons. A flag epitope-tagged B1C8 protein (corresponding to B1C8 ORF amino acids 1-820), expressed by transient transfection in HeLa cells and detected with an anti-flag antibody, precisely co-migrated with the endogenous B1C8 protein (G. Bauren, B.J.B. and P.A.S., data not shown). These analyses provided strong evidence that the first methionine residues in the predicted ORFs correspond to the sites of initiation of translation and that the ORFs represent the full length B1C8 and B4A11 proteins.

The ORFs of B1C8 and B4A11 share similar sequence features. Both contain a remarkably high content of Ser (S), Arg (R), and Pro (P) residues (summarized in Figure 1). The majority of the S, R and P residues are concentrated within several types of reiterated motifs, including multiple SR repeats, S-rich domains, and multiple R/P-rich motifs. Potential phosphorylation sites for many different kinases are located throughout the S/R/P-rich regions. The presence of multiple SR repeats in the ORFs is consistent with previous biochemical and immunological evidence indicating that B1C8 and B4A11 are SR-related

6

proteins, for example, their property of precipitation in 20mM MgCl$_2$ and reactivity with several mAbs that recognize SR-family proteins. Moreover, the presence of multiple potential phosphorylation sites is consistent with previous evidence that B1C8 and B4A11 are highly phosphorylated (Blencowe et al., 1994, 1995).

In contrast to SR-family splicing factors (see introduction), the predicted B1C8 and B4A11 ORFs lack RNA Recognition Motifs (RRMs) or RRM-homology (RRM-H) domains. Furthermore, searches of the data bases with the ORF sequence did not yield significant similarities to characterized proteins, nor to any sequence in the S. cerevisiae genome. However, candidate homologues were identified in the EST data bases for both ORFs. In each case, partial cDNA sequences were identified that share remarkable identity with N-terminal, non-repetitive, regions of the human ORFs. Sequences homologous to B1C8 are shown in Figure 1C. The candidate B1C8 homologues in Figure 2 have the following identities with the human sequence: mouse (EST 388459) =98%, C. elegans (EST CELK113G9F) =55%, A. simplex (gb:S51495)=53%, and B. malayi (EST SW3ICA1799) =59%. This analysis suggests that the predicted B1C8 ORF corresponds to a protein that is highly conserved in metazoans, but absent in yeast. Similar search results were obtained for the N-terminal 120 amino acids of the B4A11 ORF, which was high homologous to sequences in Drosophila and Nematode EST cDNAs.

Association of B1C8 with nuclear speckles and splicing complexes

To establish the relationship between the cDNA ORF and the B1C8 nuclear matrix antigen originally detected by mAb B1C8, a 153 amino acid protein fragment (ORF amino acids 7 to 160) was expressed as a GST fusion and used to immunize rabbits for the production of polyclonal antisera. The affinity purified serum (rAb B1C8), but not a corresponding pre-immune serum, specifically recognized a 160kDa protein in total nuclear extract (Figure 2A, lane 1). rAb B1C8 also specifically recognized B1C8-160kDa in a partially purified fraction, as did mAb B1C8 (data not shown). Also consistent with previous results with mAb B1C8 (Wan et al., 1994; Blencowe et al., 1994), rAb B1C8 specifically stained nuclear speckle structures in interphase cells and the spindle apparatus at metaphase (Figure 2B). However, in contrast to mAb B1C8, rAb B1C8 stained many smaller foci in addition to the predominant speckles, indicating that a form of B1C8 that is not strongly bound by mAb B1C8, is concentrated in additional nuclear foci. Identical to mAb B1C8, rAb B1C8 immunoprecipitated splicing complexes assembled on PIP85A pre-mRNA from in vitro reactions, enriching for complexes containing exon sequences but not the intron-lariat RNA (Figure 2C, lane 5). The corresponding pre-immune serum did not immunoprecipitate splicing complexes, nor did it stain nuclei (Figure 2C, lane 4; data not shown). These results provide strong evidence that the isolated cDNA encodes the 160kDa B1C8 nuclear matrix antigen and demonstrates that this antigen is a novel SR domain protein preferentially associated with nuclear "speckles" and exon-containing splicing complexes.

B1C8 is associated with B4A11 and SR family proteins

The absence of an RRM in B1C8 indicates that its association with splicing complexes could be mediated by protein-protein interactions. To identify proteins that bind to B1C8 and that may facilitate interactions with pre-mRNA, rAb B1C8 immunoprecipitates were prepared from HeLa nuclear extracts and analyzed for co-immunoprecipitated proteins. Analysis of the total co-immunoprecipitated proteins revealed only four prominent species of ~300kDa, ~220kDa, 75kDa and 40kDa (data not shown). A candidate for the 300kDa species was B4A11 since it shares similar properties to B1C8 and biochemically co-purifies with B1C8-160kDa (Blencowe et al., 1994, 1995). Since previous evidence suggests that SR domains can interact (Wu and Maniatis, 1993; Amrein et al., 1994; Kohtz et al., 1994), candidates for the 75kDa and 40kDa species were SR-family proteins of these sizes. To test

this possibility, the immunoprecipitates were also probed with mAb 104, which detects a specific SR domain phosphoepitope shared between many SR proteins (Roth et al., 1990; Zahler et al., 1992).

Monoclonal Ab B4A11 detected the B4A11 protein in the rAb B1C8 immunoprecipitate (Figures 3A and 4B, lane 3). Furthermore, mAb 104 detected proteins in the immunoprecipitate corresponding in size to SRp40 and SRp75, as well as detecting the B1C8 and B4A11 proteins (Figure 3B, lane 3). The co-immunoprecipitation of these SR-family proteins by rAb B1C8 was specific as none of the other abundant SR-family proteins, including SRp20, SRp30 and SRp55, were detected. Furthermore, none of the mAb104-reactive proteins were immunoprecipitated by the pre-immune serum (lane 4). Analysis of the supernatant fractions revealed that rAb B1C8 essentially depleted all of the B1C8 protein and efficiently co-depleted the B4A11 protein (Figure 3C and 3D, lane 2; data not shown). By contrast, only minor amounts of SRp40 and SRp75 were depleted as the majority of these proteins, as well as the other SR-family proteins, remained in the supernatant (Figure 3D, compare lanes 2 and 4). These results indicate that essentially all of the B4A11 protein in nuclear extract is bound to B1C8-160kDa and that this B1C8/B4A11 complex is specifically associated with a subfraction of SRp40 and SRp75 proteins.

Synergistic interactions between B1C8/B4A11 and SR family proteins

B1C8/B4A11 proteins could participate in splicing through interactions with SR-family proteins. To investigate this possibility, a preparation of purified B1C8/B4A11 proteins was assayed for activity in splicing in conjunction with SR-family proteins in S100 extracts. These "cytoplasmic" extracts of HeLa cells contain all factors required for splicing except SR-family proteins and are also deficient in B1C8 and B4A11 (Krainer et al., 1990, Zahler et al., 1992; data not shown). Addition of high levels of any one particular SR-family protein to an S100 extract promotes splicing activity.

Addition of purified B1C8/B4A11 to a S100 extract did not promote splicing activity or immunoprecipitation of PIP85A pre-mRNA by mAb B1C8 (Figure 4A, lanes 4-6 and 10-12). The same extract was however active for splicing upon addition of SR-family proteins (see Figure 4B below; data not shown). By contrast, addition of B1C8/B4A11 to nuclear extract stimulated splicing activity, and resulted in a corresponding increase in the level of immunoprecipitation of splicing complexes by mAb B1C8 (lanes 1-3 and 7-9). These results demonstrate that B1C8/B4A11 is functionally distinct from SR-family proteins. It binds to pre-mRNA and strongly promotes splicing activity in nuclear extracts, yet, in the absence of SR-family proteins, does not bind to pre-mRNA or restore splicing activity in S100 extracts.

To investigate if the activity of B1C8/B4A11 is mediated by interactions with the SR-family of proteins, mixing experiments were performed in S100 reactions incubated with β–globin pre-mRNA. Addition of varying amounts of B1C8/B4A11 that were not active alone in an S100 extract (Figure 4B, lane 9; data not shown), together with limiting amounts of SR-family proteins, resulted in a significant stimulation of splicing activity (compare lanes 3 and 4 with lanes 5 and 6). The maximum level of stimulation was observed when 2ug of B1C8/B4A11 proteins was added to 2ug of SR-family proteins and was not increased upon addition of higher levels of B1C8/B4A11 or SR-family proteins (compare lanes 5,6 and 7,8). Similar results were obtained with PIP85A pre-mRNA (data not shown). Also similar to PIP85A splicing reactions, B1C8/B4A11 proteins stimulated β–globin splicing when added to nuclear extract in the absence of exogenous SR-family proteins (lane 10). Additional control experiments were performed to determine if the co-stimulatory effect of the B1C8/B4A11 protein preparation was due to B1C8/B4A11 proteins and not a

8

contaminating SR protein activity. As in S100 assays, the B1C8/B4A11 preparation did not restore activity to U1 snRNP-depleted reactions incubated with β-globin pre-mRNA, that were however complemented by SR family proteins (data not shown; see Figure 7). The purified B1C8/B4A11 preparation also did not promote intron-proximal splice site utilization in a pre-mRNA containing cis-competing 5' splice sites, as did low levels of SR-family proteins (data not shown). The results in Figure 4B therefore indicate that one or both of the B1C8/B4A11 nuclear matrix proteins promotes general splicing activity in nuclear extracts and in S100 extracts supplemented with SR-family proteins.

B1C8/B4A11 is a pre-mRNA-specific splicing factor

To determine if B1C8/B4A11 proteins are essential for splicing activity at endogenous concentrations of SR-family proteins, nuclear extracts were immunodepleted of B1C8/B4A11 proteins (see Figures 3C and 3D (lanes 2) and tested for splicing activity on different pre-mRNA substrates (Figure 5). Depletion of B1C8/B4A11 prevented the first step of splicing of PIP85A pre-mRNA (Figure 5A, lane 3). Splicing activity was however fully restored upon the addition of purified B1C8/B4A11 proteins (Figure 5A, lanes 4-6). Restoration of activity in the depleted extract was specific for B1C8/B4A11 proteins since the B1C8/B4A11 protein buffer alone did not restore activity (lane 7), nor did addition of B1C8/B4A11 proteins to nuclear extracts depleted of U2 snRNP (lane 8). Addition of high levels of SR-family proteins to B1C8/B4A11 depleted reactions resulted in a partial restoration of splicing activity, yet only addition of B1C8/B4A11 proteins resulted in full restoration of activity (data not shown; see discussion). These results indicate that, at endogenous concentrations of SR-family proteins, B1C8/B4A11 proteins are required for the first step of splicing of PIP85A pre-mRNA.

Surprisingly, depletion of B1C8/B4A11 did not prevent splicing of the β–globin pre-mRNA substrate (Figure 5B, compare lanes 1 and 2). This was unexpected since antibodies to both B1C8 and B4A11 proteins immunoprecipitate β–globin splicing complexes (Blencowe et al., 1994; data not shown), and excess B1C8/B4A11 proteins stimulated β–globin pre-mRNA splicing in nuclear extract and in S100 reactions supplemented with SR proteins (Figure 4). Furthermore, it was previously found that mAbs to B1C8 and B4A11 both inhibited splicing of β–globin pre-mRNA (Blencowe et al., 1994; data not shown). However, consistent with the immunodepletion results, addition of rAb B1C8 to nuclear extracts prevented splicing of PIP85A pre-mRNA but not β–globin pre-mRNA (data not shown). The most likely interpretation of these results is that the splicing of β–globin pre-mRNA may occur by two pathways, one of which is dependent upon B1C8/B4A11. Moreover, the previously observed inhibition of β-globin splicing by mAbs B1C8 and B4A11 was probably due to one or more indirect effects of antibody binding.

Additional pre-mRNAs were tested for activity in the B1C8/B4A11 depleted nuclear extracts. A pre-mRNA from the drosophila fushi tarazu gene (Ftz), like the β–globin pre-mRNA, was spliced with similar efficiency in the depleted extract and extract mock-depleted with pre-immune serum (Figure 5C, lanes 1 and 2). In contrast, a substrate derived from the major late region of Adenovirus (Ad1), from which PIP85A pre-mRNA was derived, was spliced poorly in the depleted extract (lanes 3 and 4). Although inefficient, the level of Ad1 splicing in the B1C8/B4A11 depleted extract was higher than that of PIP85A pre-mRNA (compare lanes 2 and 3 with lanes 5 and 6). These results

9

demonstrate that specific pre-mRNAs differ with respect to their dependence on B1C8/B4A11 proteins for splicing.

Cross-intron networking interactions involving B1C8/B4A11

When compared with previous findings, the above results indicate a correlation between B1C8/B4A11 dependence and a requirement for U1 snRNP in splicing. The β–globin pre-mRNA is efficiently spliced in U1 snRNP depleted reactions supplemented with excess SR-family proteins whereas, under the same conditions, the Ad1 pre-mRNA is spliced with reduced efficiency and the PIP85A pre-mRNA is spliced poorly (Crispino et al., 1994, 1996). Ftz pre-mRNA is also spliced efficiently in the absence of U1 snRNP, but without the addition of exogenous SR-family proteins (Crispino et al., 1996). These correlative results indicate that pre-mRNAs that are not spliced in the absence of U1 snRNP, even when SR-family proteins are in excess, have a stronger requirement for B1C8/B4A11. This correlation was further strengthened by an analysis of PIP85A/β–globin chimeric pre-mRNAs, in which it was determined that sequences within the 3' half, but not 5' half of the β–globin pre-mRNA, when transferred to PIP85A, conferred efficient splicing both in the absence of B1C8/B4A11, or U1 snRNP when SR-family proteins were in excess (data not shown; Crispino et al., 1996). It is possible therefore that B1C8/B4A11 and U1 snRNP cooperate in promoting stable interactions leading to the formation of productive splicing complexes on some substrates. To investigate this possibility, B1C8/B4A11 and U1 snRNP were tested for functional interdependence.

Under regular splicing conditions, binding of B1C8/B4A11 to pre-mRNA is dependent on U1 snRNP. Immunoprecipitation of pre-mRNA by mAb B1C8 was very inefficient in splicing reactions depleted of U1 snRNP (Figure 6A, lane 8) but was only slightly reduced in reactions depleted of U2 snRNP (lane 7). The lack of pre-mRNA immunoprecipitation resulting from depletion of U1 snRNP, and the reduction resulting from depletion of U2 snRNP, was not due to non-specific losses since full levels of splicing activity and splicing complex immunoprecipitation were obtained in reactions containing an equal mix of the U1 and U2 depleted extracts (lane 9). Furthermore, these differences were not due to degradation since pre-mRNA was recovered intact in the supernatant fractions of the immunoprecipitations (data not shown). These results indicate that, under regular splicing conditions, binding of B1C8/B4A11 to pre-mRNA is greatly stimulated by the presence of U1 snRNP. In contrast, U2 snRNP is not strongly required but may stabilize B1C8/B4A11 binding.

The results in Figure 6A indicate that B1C8/B4A11 may form multiple interactions with pre-mRNA dependent on snRNP binding. To further investigate this possibility, the ability of B1C8/B4A11 to associate with separate 5' and 3' halves of PIP85A pre-mRNA was tested in the presence and absence of U1 and U2 snRNPs. mAb B1C8 immunoprecipitated both PIP85A 5' and 3' half RNAs incubated under splicing conditions in a mock-depleted nuclear extract (Figure 6B, lane 6), although at a lower efficiently than the full length PIP85A pre-mRNA (compare Figures 6A and 6B). Moreover, the level of immunoprecipitation of the 5' half RNA was approximately 3-5 fold lower than the 3' half substrate (lane 6). Depletion of U2 snRNP significantly reduced the level of immunoprecipitation of the 3' half RNA but did not affect the level of immunoprecipitation of the 5' half RNA (Figure 6A, lane 7). Similarly, depletion of U1 snRNP resulted in a reduced level of immunoprecipitation of the 5' half RNA but did not significantly alter the level of immunoprecipitation of the 3' half RNA (compare lanes 8 and 9). Immunoprecipitation of the half RNAs was restored in a reaction containing an equal mix of the U1 and U2 snRNP depleted extracts (lane 9).These results indicate that

10

B1C8/B4A11 can bind independently to the 5' half and 3' halves of PIP85A pre-mRNA and that these interactions are dependent on U1 and U2 snRNPs, respectively.

The results in Figures 6A and 6B indicate that B1C8/B4A11 could associate with one of more snRNP components in pre-mRNA splicing complexes. To determine if B1C8 interacts directly with snRNPs, the ability of rAb B1C8 to co-immunoprecipitate snRNPs in the absence of exogenous pre-mRNA was tested. rAb B1C8, but not pre-immune serum, specifically immunoprecipitated U2 snRNP from HeLa nuclear extracts (Figure 6C). Under low salt (100mM KCl) immunoprecipitation conditions, rAb B1C8 immunoprecipitated higher levels of U1, U2, U5 and U5 snRNPs than the corresponding pre-immune serum (compare lanes 2 and 3). However, only U2 snRNP was specifically immunoprecipitated at higher salt (300mM KCl) conditions (compare lanes 6 and 7). To determine if the association of B1C8/B4A11 with U1 and U2 snRNPs, under the low salt conditions, is direct, or else is mediated by indirect interactions through endogenous pre-mRNA in the nuclear extract, antisense blocking experiments were performed. Pre-incubation of extracts with an anti-U1 2'-OMe RNA oligonucleotide (complementary to U1 nucleotides 1-13; Barabino et al., 1991; see experimental procedures) that blocks the U1 snRNP-5' splice site interaction, specifically prevented immunoprecipitation of U1 but not U2 snRNP by rAb B1C8 (Figure 6D, compare lanes 3 and 7). By contrast, immunoprecipitation of U2 snRNP was not affected by blocking of U2 snRNP with a 2'-OMe RNA oligonucleotide complementary to the branch site pairing region (nucleotides 31 to 45 of U2 snRNA; Barabino et al., 1991; see experimental procedures) that prevents the U2 snRNP-pre-mRNA association (compare lanes 3 and 5). These results indicate that B1C8/B4A11 directly associates with U2 snRNP, but not U1 snRNP, in the absence of pre-mRNA.

The results in Figure 6 indicate that B1C8/B4A11 forms multiple interactions with snRNP components bound at the 5' and 3' splice sites and that, under normal splicing reactions, B1C8/B4A11 may function via a U1 snRNP-dependent assembly pathway. However, it is important to remember that endogenous levels of B1C8/B4A11 are normally rate-limiting for splicing in vitro (see Figure 4) and, consequently, may be insufficient to promote critical interactions important for splicing when U1 snRNP is absent, even when SR-family proteins are in excess. To test this possibility, purified B1C8/B4A11 was added to U1 snRNP-depleted reactions in the presence or absence of excess SR-family proteins (Figure 7). Addition of B1C8/B4A11 alone did not restore activity to U1 depleted reactions when assayed for splicing of PIP85A pre-mRNA (lanes 6 and 7). Remarkably however, addition of B1C8/B4A11 to U1 depleted reactions supplemented with excess SR proteins restored splicing activity (lanes 8-12). The highest level of activity was obtained upon addition of 1.5ug of B1C8/B4A11 and 3ug of SR-family proteins (lane 12) and was saturated upon addition of more B1C8/B4A11 (lane 13).

Addition of low levels (1ug) of SR proteins alone to the U1 depleted extract resulted in partial restoration of splicing activity to the PIP85A pre-mRNA substrate (lane 3) whereas higher levels (2-3ug) did not promote splicing activity (lanes 4 and 5). These amounts of SR-family proteins did however promote splicing of the β-globin pre-mRNA in U1 snRNP-depleted reactions (Crispino et al., 1994; data not shown). By contrast, no level of B1C8/B4A11 tested restored splicing of PIP85A or β-globin pre-mRNA in the absence of U1 snRNP. It is possible that the low level of endogenous B1C8/B4A11, which is required for splicing of PIP85A but not β-globin pre-mRNA, is titrated into non-functional complexes by addition of excess SR-family proteins in the U1 snRNP depleted reaction. These results provide further evidence for specific functional differences between B1C8/B4A11 and SR-family proteins and also indicate that not only the levels but ratios of

11

these proteins is critical for promoting efficient splicing of PIP85A pre-mRNA. Taken together with the data in Figures 4 and 6, synergistic interactions between B1C8/B4A11 and SR-family proteins may play a critical role in the splicing of specific pre-mRNAs, by both U1 dependent and independent mechanisms.

Discussion and Conclusions

B1C8/B4A11 is a novel SR protein complex that is distinct in activity from previously characterized protein splicing factors. B1C8 and B4A11 proteins contains an unusually high content of serine (S), arginine (R) and proline (P) residues and lack an RRMs. B1C8/B4A11 preferentially associates with the SR-family proteins comigrating with SRp40 and SRp75. Biochemically purified B1C8/B4A11 promotes splicing activity in nuclear extracts but, unlike SR-family proteins, not in splicing deficient S100 extracts. However, co-addition of B1C8/B4A11 and SR-family proteins to S100 cytoplasmic reactions significantly enhances splicing levels. B1C8/B4A11 normally associates with pre-mRNA via a U1 snRNP-dependent pathway and its binding to pre-mRNA is further stabilized by U2 snRNP. However, addition of B1C8/B4A11 to splicing-inactive U1 snRNP-depleted reactions supplemented with excess SR-family proteins restores activity. B1C8/B4A11 therefore cooperates with but is functionally distinct from SR-family proteins. It is therefore proposed that B1C8/B4A11 fulfills the role of a coactivator of splicing. It interacts with multiple splicing factors that directly bind to the precursor mRNA, and through these interactions, promotes splicing.

Addition of excess B1C8/B4A11 to nuclear extracts, or to S100 reactions supplemented with SR proteins, stimulated splicing activity of all substrates tested. That this general activity is rate limiting in reactions is consistent with the general immunoprecipitation of pre-mRNA splicing complexes by B1C8 and B4A11 antibodies. However, depletion of B1C8/B4A11 only inactivated the splicing of a specific subset of these pre-mRNAs. This suggests multiple pathways of interactions can result in splicing. Some pre-mRNAs are highly dependent for splicing on pathways involving interactions with B1C8/B4A11, while other pre-mRNAs utilize interactions that are less dependent or are independent of this coactivator. Consistent with complementary roles for B1C8/B4A11 and SR-family proteins on B1C8/B4A11 dependent substrates, addition of excess SR proteins partially restored splicing activity in B1C8/B4A11-depleted reactions (unpublished observations); full restoration of activity however required the addition of B1C8/B4A11. Moreover, excess SR-family proteins alone did not promote efficient splicing of PIP85A pre-mRNA in the absence of U1 snRNP; full restoration of activity also required the addition of B1C8/B4A11. These findings indicate that one or more SR-family proteins may have a parallel function with B1C8/B4A11 on some substrates, but cannot substitute on other substrates. The results indicate that B1C8/B4A11 and SR-family proteins have overlapping but non-reciprocal functions and that cooperative interactions between these factors is critical for the splicing of specific pre-mRNAs.

The above results suggest a model in which multiple interactions between U1 snRNP, B1C8/B4A11 and SR proteins, under normal splicing conditions, function in the formation of splicing complexes (see Figure 8). In this model, substrate-associated SR proteins and U1 snRNP bind B1C8/B4A11. These interactions are further stabilized by, possibly through direct association, binding of U2 snRNP to the branch site. This network of interactions could occur in conjunction with intron-bridging interactions that have recently been proposed (Abovich and Rosbash, 1997). The model is also consistent with earlier proposals for the involvement of an SR protein network involving SC35 and ASF/SF2 in the formation of bridging interactions between U1-70K protein at the 5' splice site, and the 35kDa subunit of the U2 auxiliary factor, and/or a related SR domain protein (Urp), at the 3' splice site (Wu and Maniatis, 1993; Kohtz et al., 1994; Tronchère et al., 1997). B1C8/B4A11 may contribute to such an SR interaction network, possibly specifically involving SRp40 and SRp75, that is required for the recognition of pairs of exon-intron boundaries.

13

B1C8/B4A11 is reminiscent of splicing regulators that form cooperative interactions with SR proteins on "enhancer" sequences. Splicing enhancers are typically located within exons where they stimulate the utilization of adjacent splice sites. In a prototypic example, a multiprotein complex containing the drosophila sex determination factors Tra and Tra2, both of which are SR proteins, and the RBP1 (SRp20) SR-family protein, assembles on repeats of a specific enhancer sequence within exon 4 of doublesex pre-mRNA (Tian and Maniatis, 1992,1993; Heinrichs and Baker, 1995; Lynch and Maniatis, 1996). This complex promotes the utilization of an adjacent, weak 3' splice site resulting in exon 4 inclusion (female-specific splicing pattern) vs. exon 4 exclusion (male-specific splicing pattern). Similar to B1C8, Tra is an SR protein that lacks an RRM but associates with other SR proteins that bind to RNA. However, Tra and B1C8 proteins differ in that, under normal splicing conditions, binding of the latter to pre-mRNA strongly requires U1 snRNP and is stabilized by U2 snRNP, whereas the former is not known to have this property. Thus, the B1C8/B4A11 complex appears to be functionally analogous to coactivators of transcription in that it probably promotes splicing by bridging between sequence-specific factors (SR-family proteins) and basal factors of the spliceosome (SR-family proteins and snRNP components).

A remarkable property of B1C8 is its unusually high content of serine, arginine and proline residues and multiple potential phosphorylation sites on serine, threonine and tyrosine residues. Consistent with these sequence features, B1C8 collapses from 160kDa to 120kDa upon treatment with phosphatases, indicating that it is normally highly phosphorylated in vitro. Moreover, B1C8 is detected by three different mAbs (104, NM4 and NM22) which recognize distinct phosphoepitopes shared between SR proteins and is a substrate for at least one SR-specific kinase, indicating that it is phosphorylated within its SR repeats (Blencowe et al., 1995; unpublished observations). Previous studies have demonstrated that changes in the phosphorylation state of splicing factors can have a profound effect on nuclear organization, splicing activity, and also splice site selection (reviewed by Fu, 1995; Mermoud et al., 1994a; Misteli and Spector, 1997). The abundance of potential phosphorylation sites and extensive phosphorylation of B1C8 suggests that it may be an important target of kinases and phosphatases that regulate splicing.

B1C8 and B4A11 are integral proteins of the nuclear matrix, indicating that they may function in association with this substructure (Wan et al., 1994; Blencowe et al., 1994). Previous studies support a role for the nuclear matrix in splicing. Pre-mRNAs and spliced mRNAs of specific transcripts, as well as splicing factors, have been detected in nuclear matrix preparations prepared by biochemical fractionation and/or in situ extraction (reviewed in Penman et al., 1997, Mattern et al., 1997). Moreover, it was previously demonstrated that a β–globin pre-mRNA associated with isolated nuclear matrices can be chased with rapid kinetics to spliced product upon the addition of at least one soluble factor (Zeitlin et al., 1987, 1989). The results in the present study are consistent with a previous model in which it was proposed that SR proteins associated with the nuclear matrix perform critical roles in the recognition and regulation of splice site selection (Blencowe et al., 1994). As integral matrix proteins, it is also possible that B1C8 and B4A11 function in the nuclear organization of pre-mRNA splicing. B1C8 and B4A11 are both concentrated within speckled domains in the matrix. Consistent with evidence that the majority of co-transcriptional splicing may occur outside of speckles, B1C8 is also detected in many smaller foci separated from speckle domains. The detection of specific pre-mRNAs and corresponding spliced mRNAs, at the periphery or within speckles, indicates that these structures are sites of processing of a subset of pre-mRNAs (Xing et al., 1993, 1995; Huang and Spector, 1996). It is possible that the preferential localization of these transcripts to speckle domains is influenced by their dependence on specific combinations of pre-mRNA splicing factors, including B1C8/B4A11.

14

Relevance to the Statement of Work objectives.

Task 1

The research described in this report above provides significant new insights into the structure and function of the nuclear matrix proteins B1C8 and B4A11 in splicing. The interactions of these factors with splicing complexes was investigated and characterized in some detail. This work therefore addresses the main objectives outlined in the SoW for Months 1-12. It should be noted that the proposed RNAse protection assays that were also proposed were performed. However, this method did not result in the identification of a region of protected RNA less than half the length of the pre-mRNA. The multiple series of cooperative interactions required for stable binding of B1C8/B4A11 to pre-mRNA described in this report most likely accounts for this result. An investigation of factor interactions utilizing extracts depleted of snRNP and SR protein components was, therefore, conducted to determine the nature of the interactions between B1C8/B4A11 and pre-mRNA. Very recently, we have expressed recombinant B1C8 in baculovirus. If active in reconstitution assays, mutant derivatives of B1C8 will be expressed and tested for activity. This work will be described in next year's report. It should also be noted that, given the identification of the association of B4A11 with B1C8, the completion of the cloning and sequencing of this protein was prioritized. We are currently preparing recombinant B4A11 and a polyclonal antiserum specific for this protein so as to analyze its role in splicing.

Task 2

The laboratory of Dr. S. Berget (Baylor, Texas) has recently presented experiments (at a recent Cold Spring Harbor Meeting) involving the immunological screening of different normal and cancer cell lines with anti-SR protein antibodies. Given the potential overlap of her work with experiments proposed in Task 2, I decided to focus attention on the functional characterization of B1C8/B4A11.

Additional Studies.

An unexpected development described in last year's report was the identification of interactions between a hyperphosphorylated form of the large subunit of RNA polymerase II (pol IIo) and splicing complexes. This finding suggested a mechanism for how gene expression processes may be coupled. We have also followed up on these studies. During the past year, an SR-related protein was purified that appears to be associated with both pol IIo and splicing snRNPs. Thus, a major effort during the year was dedicated to isolation of an SR-related protein that may bridge the processes of transcription and splicing. During the second year of the CDA, we will attempt to microsequence and identify cDNAs corresponding to this new SR-related protein. It is believed that this line of investigation will also contribute to the understanding of mechanisms underlying the regulation RNA processing relevant to the formation of breast cancers.

15

References

Abovich, N. and M. Rosbash (1997). Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell* **89**: 403-412.

Alzhanova-Ericsson, A. T., X. Sun, N. Visa, E. Kiseleva, T. Wurtz and B. Daneholt (1996). A protein of the SR family binds extensively to exonic Balbiani ring pre-mRNA and accompanies the RNA from gene to the nuclear pore. *Genes & Dev.* **10**: 2881-2893.

Amrein, H., M. L. Hedley and T. Maniatis (1994). The role of specific protein-RNA interactions and protein-protein interactions in positive and negative control of pre-mRNA splicing by *transformer* 2. *Cell* **76**: 735-746.

Barabino, S. M. L., B. J. Blencowe, U. Ryder, B. S. Sproat and A. I. Lamond (1990). Targeted snRNP depletion reveals an additional role for mammalian U1 snRNP in spliceosome assembly. *Cell* **63**: 293-302.

Berget, S. (1995). Exon recognition in vertebrate splicing. *J. Biol. Chem.* **270**: 2411-2414.

Black, D. (1995). Finding splice sites within a wilderness of RNA. *RNA* **1**: 763-771.

Blencowe, B. J., R. Issner, J. Kim, P. McCaw and P. A. Sharp (1995). New proteins related to the Ser-Arg family of splicing factors. *RNA* **1**: 852-865.

Blencowe, B. J., J. A. Nickerson, R. Issner, S. Penman and P. A. Sharp (1994). Association of nuclear matrix antigens with exon-containing splicing complexes. *J. Cell Biol.* **127**: 593-607.

Blencowe, B. J., B. S. Sproat, U. Ryder, S. Barabino and A. I. Lamond (1989). Antisense probing of the human U4/U6 snRNP with biotinylated 2'-OMe RNA oligonucleotides. *Cell* **59**: 531-539.

Bregman, D., L. Du, S. van der Zee and S. Warren (1995). Transcription-dependent redistribution of the large subunit of RNA polymerase II to discrete nuclear domains. *J Cell Biol* **129**: 287-298.

Brendel, V., P. Bucher, I. Nourbakhsh, B. E. Blaisdell and S. Karlin (1992). Methods and algorithms for statistical analysis of protein sequences. *Proc. Natl. Acad. Sci. USA.* **89**: 2002-2006.

Caceres, J. F., T. Mistelli, G. R. Screaton, D. L. Spector and A. R. Krainer (1997). Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J. Cell Biol.* **138**: 225-238.

Carter, K. C., K. L. Taneja and J. B. Lawrence (1991). Discrete nuclear domains of poly(A) RNA and their relationship to the functional organization of the nucleus. *J. Cell Biol.* **115**: 1191-1202.

Chabot, B. (1996). Directing alternative splicing: cast and scenarios. *Trends in Genetics* **12**: 472-478.

16

Colwill, K., T. Pawson, B. Andrews, J. Prasad, J. Manley, J. Bell and P. Duncan (1996). The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. *EMBO J* **15**: 265-275.

Crispino, J. D., B. J. Blencowe and P. A. Sharp (1994). Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science* **265**: 1866-1869.

Crispino, J. D., J. E. Mermoud, A. I. Lamond and P. A. Sharp (1996). Cis-acting elements distinct from the 5' splice site promote U1-independent pre-mRNA splicing. *RNA* **2**: 664-673.

Du, L. and S. Warren (1997). A functional interaction between the carboxy-terminal domain of RNA polymerase II and pre-mRNA splicing. *J Cell Biol* **136**: 5-18.

Fakan, S. and E. Puvion (1980). The ultrastructural visulazation of nuclear and extranucleolar RNA synthesis and distribution. *Int. Rev. Cyt.* **65**: 255.

Fay, F., K. Taneja, S. Shenoy, L. Lifshitz and R. Singer (1997). Quantitative digital analysis of diffuse and concentrated nuclear distributions of nascent transcripts, SC35 and poly(A). *Exp Cell Res* **231**: 27-37.

Fu, X.-D. (1995). The superfamily of arginine/serine-rich splicing factors. *RNA* **1**: 663-680.

Fu, X.-D. and T. Maniatis (1990). Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature* **343**: 437-441.

Gui, J.-F., W. S. Lane and X.-D. Fu (1994). A serine kinase regulates intracellular localization of splicing factors in the cell cycle. *Nature* **369**: 678-682.

Harlow, E. and D. Lane (1988). Antibodies: a laboratory manual. Cold Spring Harbor Laboratory Press. New York.

Hedley, M., H. Amrein and T. Maniatis (1995). An amino acid sequence motif sufficient for subnuclear localization of an arginine/serine-rich splicing factor. *Proc Natl Acad Sci U S A* **92**: 11524-11528.

Heinrichs, V. and B. S. Baker (1995). The Drosophila SR protein RBP1 contributes to the regulation of doublesex alternative splicing by recognizing RBP1 RNA target sequences. *EMBO J* **14**: 3987-4000.

Huang, S., T. Deerinck, M. Ellisman and D. Spector (1994). In vivo analysis of the stability and transport of nuclear poly(A)+ RNA. *J Cell Biol* **126**: 877-899.

Huang, S. and D. Spector (1996). Intron-dependent recruitment of pre-mRNA splicing factors to sites of transcription. *J Cell Biol* **133**: 719-732.

Jackson, D. A. and P. R. Cook (1985). Transcription occurs at the nucleoskeleton. *EMBO (Eur. Mol. Biol. Organ.) J.* **4**: 919-925.

Kim, E., L. Du, D. Bregman and S. Warren (1997). Splicing factors associate with hyperphosphorylated RNA polymerase II in the absence of pre-mRNA. *J Cell Biol* **136**: 19-28.

17

Kohtz, J. D., S. F. Jamison, C. L. Will, P. Zuo, R. Luhrmann, M. A. Garcia-Blanco and J. L. Manley (1994). Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature* **368**: 119-124.

Krainer, A. R., G. C. Conway and D. Kozak (1990). Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes Dev.* **4**: 1158-1171.

Kramer, A. (1996). The structure and function of proteins involved in mammalian pre-mRNA splicing. *1996* **65**: 367-409.

Lamond, A. I. and M. Carmo-Fonseca (1993). The coiled body. *Trends Cell Biol.* **3**: 198-204.

Lawrence, J., K. Carter and X. Xing (1993). Probing functional organization within the nucleus: is genome structure integrated with RNA metabolism? *Cold Spring Harbor Symp Quan Biol* **58**: 807-818.

Li, H. and P. M. Bingham (1991). Arginine/serine-rich domains of the $su(w^a)$ and *tra* RNA processing regulators target proteins to a subnuclear compartment implicated in splicing. *Cell* **67**: 335-342.

Lynch, K. and T. Maniatis (1996). Assembly of specific SR protein complexes on distinct regulatory elements of the *Drosphila doublesex* splicing enhancer. *Genes & Dev* **10**: 2089-2101.

Manley, J. and R. Tacke (1996). SR proteins and splicing control. *Genes & Dev* **10**: 1569-1579.

Mattern, K., R. v. Driel and L. d. Jong (1997). Composition and structure of the nuclear matrix. Nuclear Structure and Gene Expression. New York, Academic Press. 87-110.

McCracken, S., N. Fong, K. Yankulov, S. Ballantyne, G. Pan, J. Greenblatt, S. Patterson, M. Wickens and D. L. Bentley (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385**: 357-361.

Meier, U. and G. Blobel (1992). Nopp140 shuttles on tracks between nucleus and cytoplasm. *Cell* **70**: 127-138.

Mermoud, J., C. Calvio and A. I. Lamond (1994a). Uncovering the role of Ser/Thr protein phosphorylation in nuclear pre-mRNA splicing. *Adv Prot Phosphatases* **8**: 99-118.

Mermoud, J., P. T. W. Cohen and A. I. Lamond (1994b). Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism. *EMBO J.* **13**: 5679-5688.

Misteli, T., J. F. Caceres and D. L. Spector (1997). The dynamics of a pre-mRNA splicing factor in living cells. *Nature* **387**: 523-527.

Misteli, T. and D. Spector (1996). Serine/threonine phosphatase 1 modulates the subnuclear distribution of pre-mRNA splicing factors. *Mol Biol Cell* **7**: 1559-1572.

Misteli, T. and D. Spector (1997). Protein phosphorylation and the nuclear organization of pre-mRNA splicing. *Trends in Cell Biol* **7**: 135-138.

Mortillaro, M., B. Blencowe, X. Wei, H. Nakayasu, L. Du, S. Warren, P. Sharp and R. Bezerkney (1996). A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. *Proc Natl Acad Sci U S A* **93**: 8253-8257.

Nickerson, J. A., G. Krockmalnic, K. M. Wan, C. D. Turner and S. Penman (1992). A normally masked nuclear matrix antigen that appears at mitosis on cytoskeleton filaments adjoining chromosomes, centrioles and midbodies. *J. Cell Biol.* **116**: 977-987.

Penman, S., B. Blencowe and J. Nickerson (1997). The nuclear matrix: past and present. Nuclear Structure and Gene Expression. New York, Academic Press. 3-24.

Reed, R. (1996). Initial splice-site recognition and pairing during pre-mRNA splicing. *Curr Opin Genet Dev* **6**: 215-220.

Roscigno, R. and M. Garcia-Blanco (1995). SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. *RNA* **1**: 692-706.

Roth, M. B., C. Murphy and J. G. Gall (1990). A monoclonal antibody that recognizes a phosphorylated epitope stains lampbrush chromosome loops and small granules in the amphibian germinal vesicle. *J. Cell. Biol.* **111**: 2217-2223.

Sharp, P. (1994). Split genes and RNA splicing. *Cell* **6**: 805-815.

Smith, H. C., S. G. Harris, M. Zillmann and S. M. Berget (1989). Evidence that a nuclear matrix protein participates in pre-mRNA splicing. *Exp. Cell Res.* **182**: 521-533.

Spector, D. (1993). Macromolecular domains within the cell nucleus. *Ann Rev Cell Biol* **9**: 265-315.

Spector, D. L., X.-D. Fu and T. Maniatis (1991). Associations between distinct pre-mRNA splicing components and the cell nucleus. *EMBO J.* **10**: 3467-3481.

Spector, D. L., W. H. Schrier and H. Busch (1983). Immunoelectron microscopic localization of snRNPs. *Biol. Cell.* **49**: 1-10.

Steinmetz (1997). Pre-mRNA processing and the CTD of RNA polymerase II: the tail that wags the dog? *Cell* **89**: 491-494.

Tian, M. and T. Maniatis (1992). Positive control of pre-mRNA splicing in vitro. *Science* **256**: 237-240.

Tian, M. and T. Maniatis (1993). A splicing enhancer complex controls alternative splicing of *doublesex* pre-mRNA. *Cell* **74**: 105-114.

Tronchere, H., J. Wang and X.-D. Fu (1997). A protein related to splicing factor U2AF35 that interacts with U2AF65 and SR proteins in splicing of pre-mRNA. *Nature* **388**: 397-400.

Valcarcel, J. and M. R. Green (1996). The SR protein family: pleiotropic functions in pre-mRNA splicing. *Trends in Biochem. Sci.* **21**: 296-301.

Verheijen, R., H. Kuijpers, P. Vooijs, W. v. Venrooij and F. Ramaekers (1986). Distribution of the 70k U1 RNA-associated protein during interphase and mitosis:

19

correlation with other U RNP particles and proteins of the nuclear matrix. *J. Cell Sci.* **86**: 173-190.

Vincent, M., P. Lauriault, M. Dubois, S. Lavoie, O. Bensaude and B. Chabot (1996). The nuclear matrix protein p255 is a highly phosphorylated form of RNA polymerase II largest subunit which associates with spliceosomes. *Nucleic Acids Res* **24**: 4649-4652.

Vogelstein, B. and B. F. Hunt (1982). A subset of small nuclear ribonucleoprotein particle antigens is a component of the nuclear matrix. *Biochem. Biophys. Res. Commun.* **105**: 1224-1232.

Wan, K., J. A. Nickerson, G. Krockmalnic and S. Penman (1994). The B1C8 protein is in the dense assemblies of the nuclear matrix and relocates to the spindle and pericentriolar filaments at mitosis. *Proc. Natl. Acad. Sci. USA* **91**: 594-598.

Wansink, D. G., W. Schul, I. v. d. Kraan, B. v. Steensel, R. v. Driel and L. d. Jong (1993). Fluorescent labeling of Nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. *J. Cell Biol.* **122**: 283-293.

Wu, J. Y. and T. Maniatis (1993). Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* **75**: 1061-1070.

Xiao, S. H. and J. L. Manley (1997). Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes and Dev.* **11**: 334-344.

Xing, Y., C. Johnson, P. M. Jr, J. McNeil and J. Lawrence (1995). Nonrandom gene organization: structural arrangements of specific pre-mRNA transcription and splicing with SC-35 domains. *J Cell Biol* **131**: 1635-1647.

Xing, Y., C. V. Johnson, P. R. Dobner and J. B. Lawrence (1993). Higher level organization of individual gene transcription and splicing. *Science* **259**: 1326-1330.

Yuryev, A., M. Patturajan, Y. Litingtung, R. Joshi, C. Gentile, M. Gebara and J. Corden (1996). The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *Proc. Natl. Acad. Sci. USA.* **14**: 6975-6980.

Zahler, A. M., W. S. Lone, J. A. Stalk and M. B. Roth (1992). SR proteins: A conserved family of pre-mRNA splicing factors. *Genes Dev.:* 837-847.

Zeitlin, S., A. Parent, S. Silverstein and A. Efstratiatdis (1987). Pre-mRNA splicing and the nuclear matrix. *Mol. Cell. Biol.* **7**: 111-120.

Zeitlin, S., R. C. Wilson and A. Efstratiatdis (1989). Autonomous splicing and complementation of in vivo assembled spliceosomes. *J. Cell Biol.* **108**: 765-777.

20

Figure Legends

Figure 1. Isolation of cDNAs encoding the B1C8 and B4A11 nuclear matrix proteins.

B1C8 and B4A11 were purified by conventional chromatography, microsequenced and the corresponding cDNAs were isolated. The ORF structures are shown. Both proteins contain highly conserved and unique N-terminal sequences (white boxes). The remaining sequences in both proteins are highly enriched in arginine (R), serine (S) and proline (P) amino acid residues (light gray shading). Approximate positions of SR dipeptides are indicated by vertical bars. Domains containing extensive stretches of SR repeats are indicated as black boxes. Poly-S domains are indicated as dark gray boxes. Neither protein contains an RNA Recognition Motif, as found in proteins of the SR-family.

B. A northern blot containing poly(A)+ from different human tissues was probed with a cDNA fragment corresponding to amino acids 1-473 of the B1C8 ORF. The same blot was reprobed with a β–actin probe (lower panel) to control for loading differences (note two forms of β–actin are expressed in heart and skeletal muscle).

C. Multiple alignment of sequences homologous to the N-terminal 150 amino acids of human B1C8. Partial cDNA ORF sequences from mouse (M. musculus) and three namatode species (C. elegans, A. simplex and B. malayi) identified by BLAST searches were aligned using the Clustal algorithm. Residues identical to the majority sequence (top line) are boxed. N-terminal residues not present in the EST cDNAs are indicated by a dash;. undefined residues are indicated by an X or blank space.

Figure 2. Association of B1C8 with exon-containing splicing complexes, interphase nuclear speckles and the mitotic spindle apparatus.

A. Total HeLa nuclear extract was separated on an SDS polyacrylamide gel and immunoblotted with an affinity purified polyclonal antiserum raised to a GST-B1C8 fusion protein containing B1C8 ORF amino acids 7-160 (rAb B1C8, lane 1), and the corresponding pre-immune serum (lane 2).

B. Interphase (top row) and metaphase (bottom row) human CaSki cells were double-immunolabeled with mAb B1C8 (red) and rAb B1C8 (green). The images were superimposed and sites of overlap between the immunostaining patterns were pseudocolored in yellow. Bar= 5um.

C. Immunoprecipitation of splicing complexes from splicing reactions incubated for 40 min with PIP85A pre-mRNA. RNA recovered following immunoprecipitation (lanes 2-5) and RNA recovered directly from a parallel splicing reaction (lane 1), was separated on a denaturing polyacrylamide gel. RNA recovered directly from the total reaction in lane 1 represents 50% of the amount of the RNA recovered from each immunoprecipitation. Immunoprecipitations were performed with a non-specific control rAb (=rabbit anti-mouse, lane 2), mAb B1C8 (lane 3), rAb pre-immune serum (lane 4) and rAb B1C8 (lane 5).

Figure 3. A complex of B1C8 and the 300kDa B4A11 nuclear matrix protein associates with the SR-family proteins SRp40 and SRp75.

A,B. Proteins immunoprecipitated from HeLa nuclear extract by rAb B1C8 were separated by SDS-PAGE, blotted and probed with mAb B4A11 (panel A), followed by mAb104 (panel B). Protein immunoprecipitated from approximately ~0.5mg nuclear extract was loaded in lanes 2-4 and ~0.15mg total nuclear extract was loaded in lane 1. Immunoprecipitations were performed with protein A beads alone (lane 2), rAb B1C8 (lane

3) and pre-immune serum (lane 4). Sizes of SR proteins, B1C8 and B4A11 are indicated in kDa. HC= immunoglobulin heavy chain.

C,D. Supernantant fractions from the immunoprecipitations in A,B were separated and immunoblotted with mAbs B4A11 and 104, as in A,B. Total nuclear extract proteins are shown in lanes 1, and proteins recovered from the immunoprecipitation supernatant fractions prepared with rAb B1C8, protein A beads alone and pre-immune serum, are shown in lanes 2, 3 and 4, respectively. ~0.15mg of nuclear extract and of each supernatant fraction was loaded per lane. A sample of purified SR-family proteins (~4ug) was separated as a marker in lane 5.

Figure 4. Pre-mRNA binding and promotion of splicing by B1C8/B4A11 requires SR-family proteins.

A. Immunoprecipitations were performed from splicing reactions incubated with PIP85A pre-mRNA for 40 min. The splicing reactions contained nuclear (lanes 1-3, 7-9) or S100 cytoplasmic (lanes 4-6, 10-12) extracts, with (lanes 2,3,5,6,8,9,11,12) or without (lanes 1,4,7,10) purified B1C8/B4A11 proteins added. 50% of the total RNA recovered directly from each splicing reaction (Totals) was loaded in lanes 1-6), whereas all of the RNA recovered following immunoprecipitation with mAb B1C8 (Pels) was loaded in lanes 7-12.

B. S100 splicing reactions (lanes 2-9) containing β–globin pre-mRNA and varying amounts of B1C8/B4A11 and SR-family proteins, as indicated, were incubated for 1 hour. Control splicing reactions incubated in parallel contained nuclear extract (lane 1) or S100 extract (lane 2), without added proteins, and a reaction containing nuclear extract and 4ug B1C8/B4A11 (lane 10). The asterisk indicates an RNA fragment protected from endogenous nuclease activity.

Figure 5. B1C8/B4A11 is required for splicing of specific pre-mRNAs.

A. Depletion of B1C8/B4A11 blocks the first step of splicing of PIP85A pre-mRNA. Splicing reactions containing nuclear extract depleted of B1C8/B4A11 proteins (see Figure 3 C,D lane 2) were incubated for 1 hour with PIP85A pre-mRNA in the presence (lanes 4-6), or absence (lanes 3,7), of purified B1C8/B4A11 proteins. Control splicing reactions contained regular nuclear extract (lane 1), nuclear extract mock-depleted with pre-immune serum (lane 2; see Figure 3C,D lane 4), B1C8/B4A11-depleted extract plus protein buffer (lane 7), and U2 snRNP depleted nuclear extract plus B1C8/B4A11 proteins (lanes 8). Splicing reactions were performed in 30ul reactions containing a range of 2-8 ug of B1C8/B4A11 proteins in lanes 4-6 and 8ug of B1C8/B4A11 proteins in lane 8.

B. Depletion of B1C8/B4A11 does not block splicing of a β–globin pre-mRNA. β–globin splicing reactions incubated for 90 min contained nuclear extract depleted of B1C8/B4A11 proteins (lane 1), nuclear extract mock-depleted with pre-immune serum (lane 2), or a non-depleted nuclear extract (lane 3).

C. Comparision of the splicing activity of different pre-mRNAs in mock depleted (lanes 1,3,5) vs. B1C8/B4A11 depleted (lanes 2,4,6) nuclear extracts. Reactions were incubated for 1hr with the drosophila fushi tarazu (ftz, lanes 1,2); Adeno major late (Ad1, lanes 3,4) and PIP85A (PIP, lanes 5 and 6) pre-mRNAs.

Figure 6. snRNP dependence for binding of B1C8/B4A11 to pre-mRNA.

A. Splicing complexes were immunoprecipitated with mAb B1C8 from snRNP depleted reactions (lanes 6-9). Splicing reactions incubated for 40 min with PIP85A pre-mRNA contained mock-depleted (lane 1), U2 -depleted (lane 2), U1 -depleted (lane 3), or an equal mix of U1 and U2 -depleted nuclear extracts (lane 4). RNA recovered directly from the splicing reactions in lanes 1-4 represents 50% of the total sample, whereas all of the

recovered from each immunoprecipitation was loaded in lanes 5-9. A control immunoprecipitation (lane 5) was performed with a non-specific antibody from a splicing reaction containing mock-depleted nuclear extract, as in lane 1.

B. Complexes assembled on separate 5' and 3' halves of PIP85A pre-mRNA were immunoprecipitated with mAb B1C8 in the presence or absence of individual snRNPs (lanes 6-9). Equal amounts of the two half RNAs were incubated under splicing conditions for 40 min in mock-depleted (lanes 1 and 2), U2 -depleted (lane 3), U1 -depleted (lane 4), or an equal mix of U1 and U2 -depleted nuclear extracts (lane 5). RNA recovered directly from the splicing reactions in lanes 1-5 represents 50% of the total sample, whereas all of the recovered from each immunoprecipitation was loaded in lanes 5-9. A control immunoprecipitation (lane 5) was performed with a non-specific antibody from a reaction containing mock-depleted nuclear (lane 1).

C. rAb B1C8 co-immunoprecipitates U2 snRNP in the absence of exogenous pre-mRNA. Nuclear extract was incubated under splicing conditions for 15 min and added to protein A beads pre-coupled to rAb B1C8. Immunoprecipitation was performed in the presence of a 100mM KCl wash buffer, and was followed by bead washes using buffers containing different concentrations of KCl, as indicated. RNA recovered from the washed beads was separated in a 10% denaturing polyacrylamide gel and analyzed by Northern blotting using snRNA-specific riboprobes. Immunoprecipitations were performed with pre-immune serum (lanes 2,4,6) or rAb B1C8 (lanes 3,5,7). A sample of total HeLa RNA was loaded as a marker in lane 1, corresponding to approximately 30% of the total input HeLa RNA (as extract) in each immunoprecipitation lane. All of the RNA recovered following immunoprecipitation was loaded in lanes 2-7.

D. Analysis of snRNP immunoprecipitation by rAb B1C8 from antisense "masked" nuclear extracts. Nuclear extracts prior incubated with an antisense 2'-OMe RNA oligonucleotides to U1 snRNP (lanes 6,7) or to U2 snRNP (lane 4,5) (see experimental procedures) were incubated under splicing conditions for 15 min and added to protein A beads pre-coupled to rAb B1C8. Immunoprecipitation was performed in the presence of a 100mM KCl wash buffer, and was followed by washes of the beads using a 150mM KCl buffer. Immunoprecipitations were performed with pre-immune serum (lanes 2,4,6) or rAb B1C8 (lanes 3,5,7). RNA was analyzed and loaded as described in Figure 6C.

Figure 7. B1C8/B4A11 coactivates splicing in U1 snRNP depleted reactions supplemented with SR-family proteins.

B1C8/B4A11 proteins were added alone (lanes 6,7), or in combination with SR-family proteins (lanes 8-13), to splicing reactions depleted of U1 snRNP. U1-depleted reactions containing excess SR proteins and no B1C8/B4A11 are shown in lanes 3-5. Control reactions containing mock depleted extract or U1-depleted extract with no added proteins are shown in lanes 1 and 2, respectively. The splicing reactions were incubated for 60 min prior to recovery of RNA.

Figure 8. Model for interactions between the B1C8/B4A11 nuclear matrix protein complex and other factors required for the formation of early splicing complexes.

A B1C8/B4A11 complex promotes splice-site pairing and splicing through multiple cooperative interactions with factors bound to pre-mRNA, including SR-family proteins, U1 and U2 snRNPs. A role for the SR domains of these factors in mediating an intron-wide network of interactions is emphasized. BS =branch site, (Py)n =polypyrimidine tract, SS =splice site.

23

Publications and Abstracts

A. Refereed Publications

1. Crispino, J.D., Blencowe, B.J. and Sharp, P.A. (1994). Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science.* 265, 1866-1869.

2. Blencowe, B.J., Nickerson, J.A., Issner, R., Penman, S. and Sharp, P.A. (1994). Association of nuclear matrix antigens with exon-containing splicing complexes. *J. Cell Biol.* 127, 593-607.

3. Blencowe, B.J., Issner, R., Kim J., McCaw, P. and Sharp, P.A. (1995). New proteins related to the Ser-Arg family of splicing factors. *RNA* 1, 852-865.

4. Nickerson, J.A. Blencowe, B.J. and S. Penman (1995). The architectural organization of nuclear metabolism. *Int. Rev. Cyt.* 162, 67-124.

5. Mortillaro, M., Blencowe, B.J., Wei, X., Nakayasu, H., Du, L., Warren, S., Sharp, P.A., and Berezney, R. (1996). A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. *Proc. Natl. Acad Sci. USA.* 93, 8253-8257.

6. Penman, S. Blencowe, B.J. and Nickerson, J.A. (1997). The nuclear matrix: past and present. In: Nuclear structure and gene expression. Stein, G. (Ed.) Academic Press. 3-24.

7. Blencowe B.J., Issner, R., Nickerson, J.A. and Sharp, P.A. A coactivator of pre-mRNA splicing. Submitted.

B. Meeting Abstracts

1. <u>Blencowe, B.J.</u>, Nickerson, J.A., Issner, R., Penman, S. and Sharp, P.A. (1994). Association of nuclear matrix antigens with exon-containing splicing complexes. RNA Society Meeting, Madison, WI.

2. Crispino, J.D., Blencowe, B.J. and Sharp, P.A. (1994). Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. RNA Society Meeting, Madison, WI.

3. <u>Blencowe, B.J.</u>, Nickerson, J.A., Issner, R., Penman, S. and Sharp, P.A. (1995). Nuclear structure- function relationships in pre-mRNA processing: cloning and characterization of the B1C8 nuclear matrix splicing factor. Keystone Symposium, Hilton Head, SC. *J. Cell. Biochem. Suppl.* 21B.

4. <u>Blencowe, B.J.</u>, Mortillaro, M., Wei, X., Nakayasu, H., Du, L., Warren, S., Sharp, P.A., and Berezney, R. (1996). A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. Cambridge Symposium, Bolton Valley, VT.

24

5. Blencowe B.J., Issner, R., Nickerson, J.A. and Sharp, P.A. (1997). A complex of novel nuclear matrix proteins functions as a splicing coactivator. Eukaryotic RNA Processing Meeting, Cold Spring Harbor, NY.

6. Blencowe B.J., Issner, R., Nickerson, J.A. and Sharp, P.A. A coactivator of pre-mRNA splicing. USDoD Breast Cancer Research Program Meeting. (1997). To be held in Washington, DC.

B1C8

N ▭▮▮▮▮▮▮▮▮▮▮ C (820 aa's)

B4A11

N ▭▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮ C (2245 aa's)

| SR dipeptide
■ SR dipeptide-rich domain
▨ S/R/P-rich
■ poly-S domain

Figure 1A

Figure 1B

```
     - - A G F F R G T S A E Q D N R F S D K E K K L L K Q M K F E E A L E K K I D M   Majority
                  10                  20                  30                  40

  1  M D A G F F R G T S A E Q D N R F S N K Q K K L L K Q L K F A E C L E K K V D M   H. sapiens
  1  - - A G F F R G T S A X X D N R F S N N K Q K K L L K Q L K F A E C L E K K V D M   M. musculus
  1  - - - - - - R G A N S E Q D G R F S D K E K K L L K Q T M K F E P Q L E Q K I D L   C. elegans
  1  - - - - - - - - - - - F T D K E R K L L K Q M R F E D A L E D K I C M   A. simplex
  1  - D A G F Y R G T S E Q D S R F T T D K E R K L L K Q M R F E E A L D E K I C M   B. malayi

     S R V N L D V I K P W I T A R V N E I L G F E D D V V I E Y I F S Q L E E K N L   Majority
                  50                  60                  70                  80

 41  S K V N L E V I K P W I T K R V T E I L G F E D D V V I E F I F N Q L E V K N P   H. sapiens
 39  S K V N L E V I K P W I T K R V T E I L G F E D D V V I E F I F N Q L   M. musculus
     N R C N M D V I K P W I T A R V N D I L G M E D D V V E Y I L S Q I D D K N L   C. elegans
 35  S R I N L D V L K P W I T A K V N E I L G M E D D V L I E Y V V F S Q L E E K T I   A. simplex
 40  D R V N L D V L K P W I T A K L N D I L G I E D D V I E Y V V F S Q L E E K S L   B. malayi

     X P K L M Q I N L T G F L N A R R A R E F M G E L W - L L L E A Q - S E D G I P   Majority
                  90                 100                 110                 120

 81  D S K M M Q I N L T G F L N G K N A R E F M G E L W P L L L S A Q E N I A G I P   H. sapiens
 73                                                                                  M. musculus
 75  T P K L L Q I N V T G F L N A R R A R E F V G D L W N L L I E A N A S E D G I P   C. elegans
 65  N P K L C K S I L T G F L N A R R A A E F L G E L W T M L L E A Q N S E D G I P   A. simplex
 80  - - K L M Q I N L T G F L N A R R A R E F M G   B. malayi

     A S L V                                                                          Majority

121  S A F L   H. sapiens
 73           M. musculus
115  A S L V   C. elegans
105  A S L V   A. simplex
100           B. malayi
```
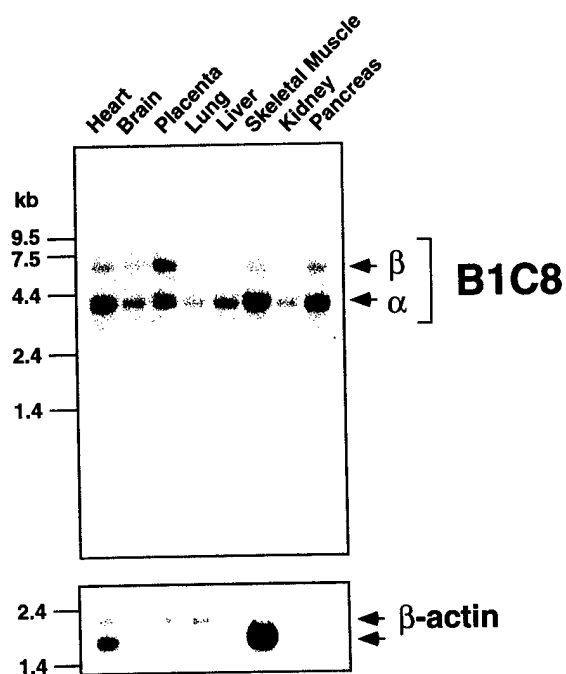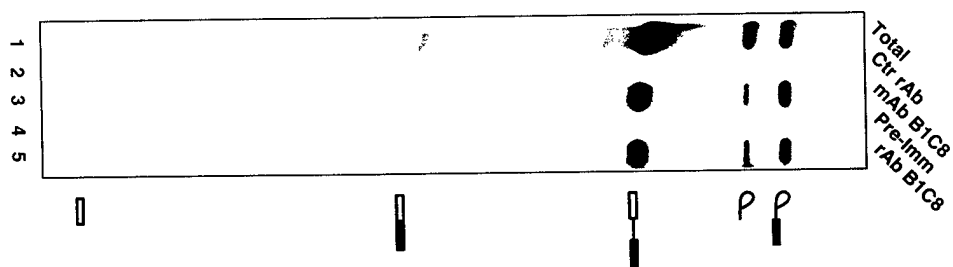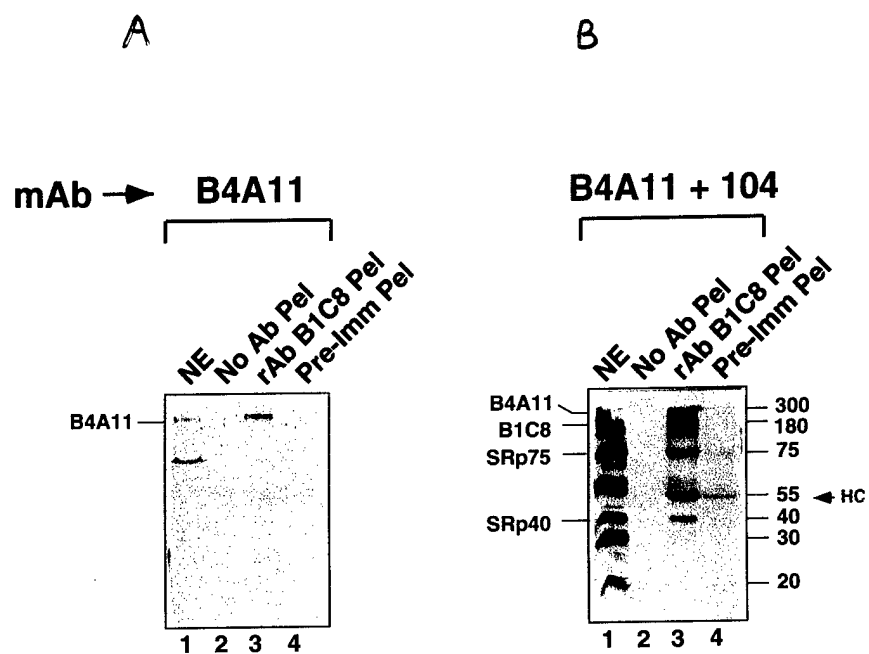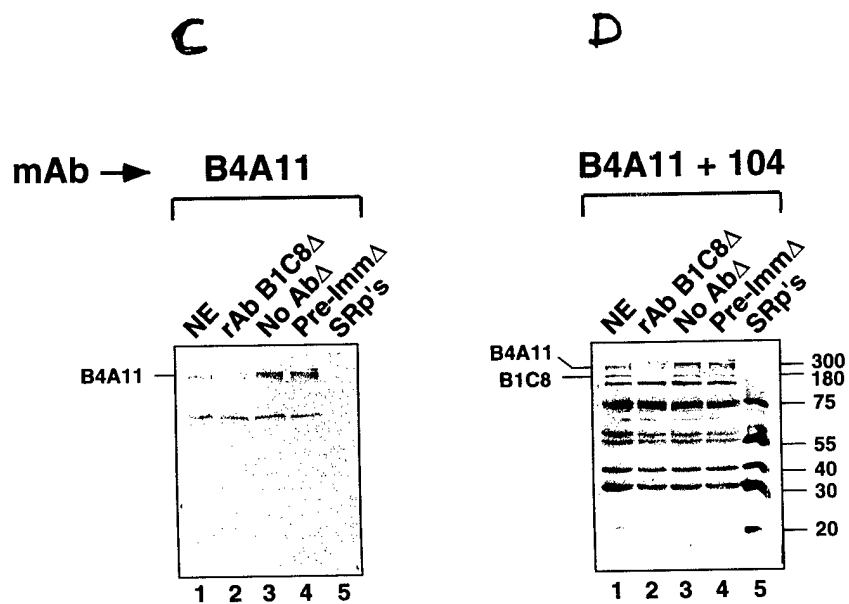
Figure 1C

Figure 2A
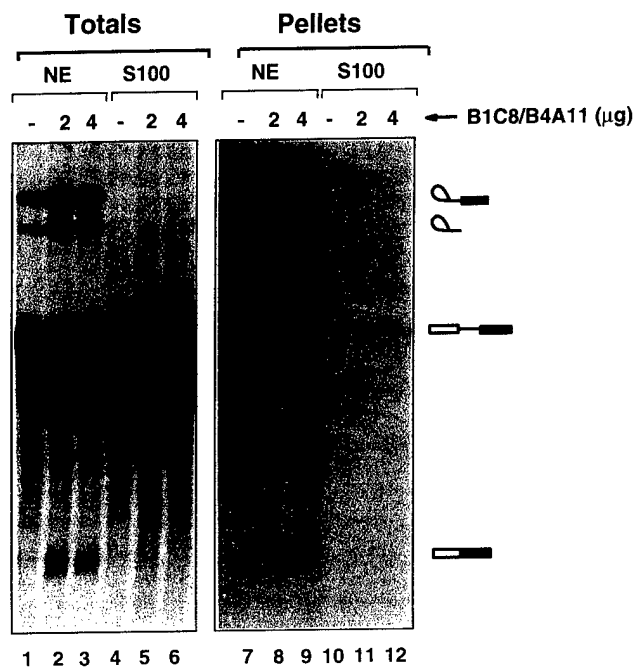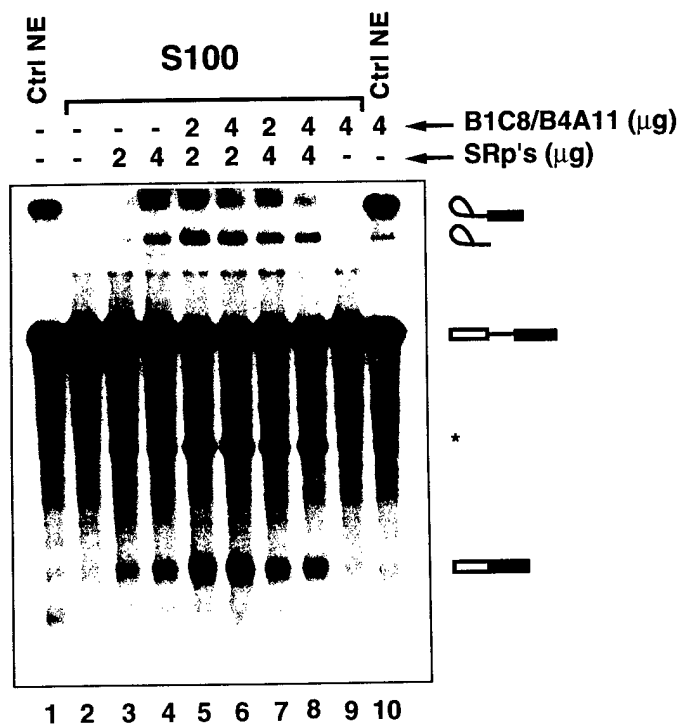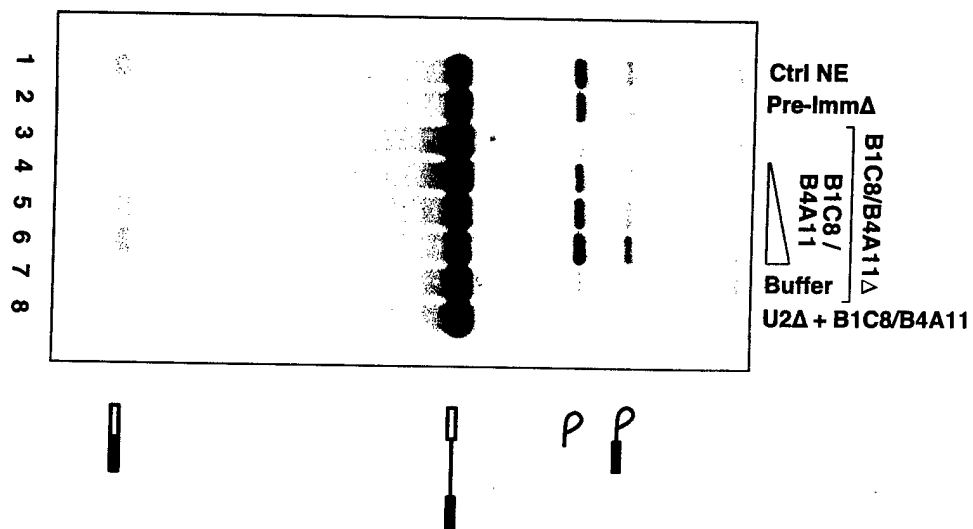
mAb   rAb   Both
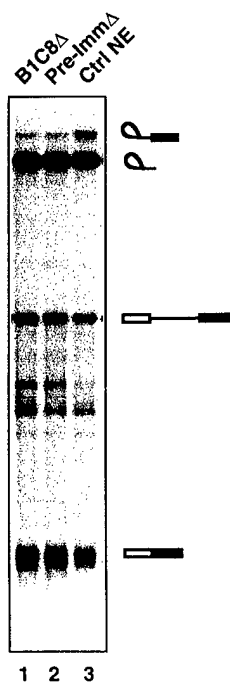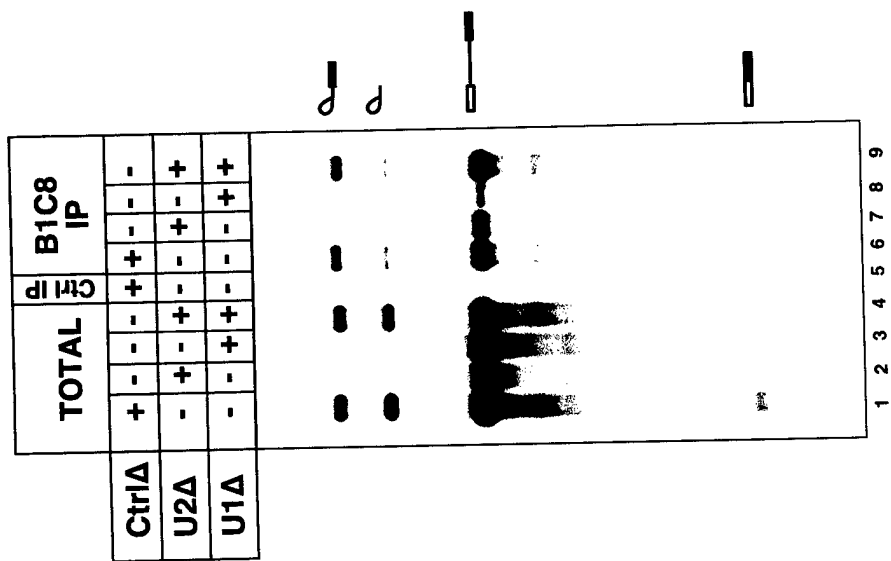

Figure 2B

Figure 2C

Figure 3

Figure 4A

Figure 4B

Figure 5A

Figure 5B

Figure 6A

Figure 5C

Figure 6B

	TOTAL					Ctrl IP	B1C8 IP			
CtrlΔ	+	+	-	-	-	+	+	-	-	-
U2Δ	-	-	+	-	+	-	-	+	-	+
U1Δ	-	-	-	+	+	-	-	-	+	+

1 2 3 4 5   6 7 8 9

Figure 6B

Figure 6C

	100mM		150mM		300mM		[KCl]
Total	-	+	-	+	-	+	rAb B1C8 IP
	+	-	+	-	+	-	Pre-Imm IP

U2 —
U1 —
U4 —
U5 —
U6 —

1 2 3 4 5 6 7

Figure 6C (1)

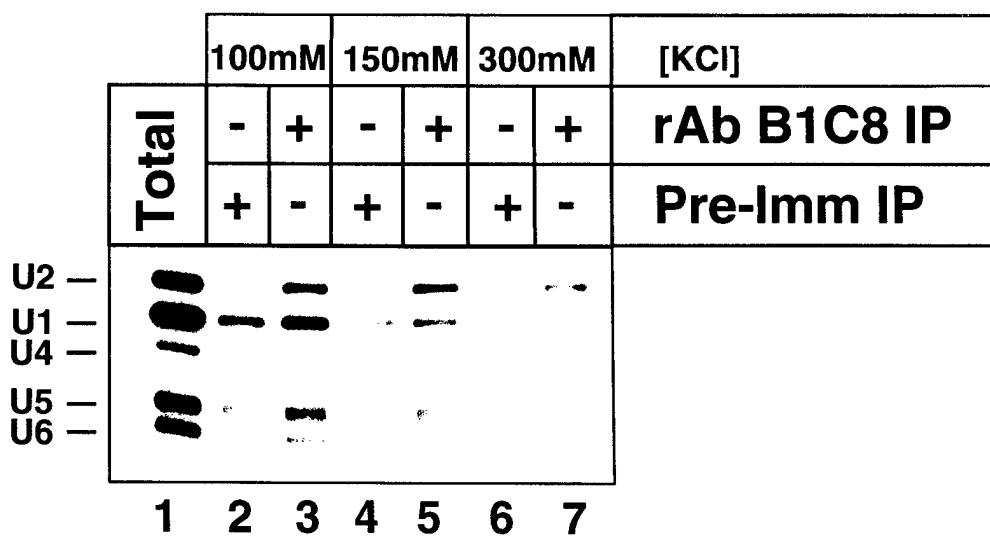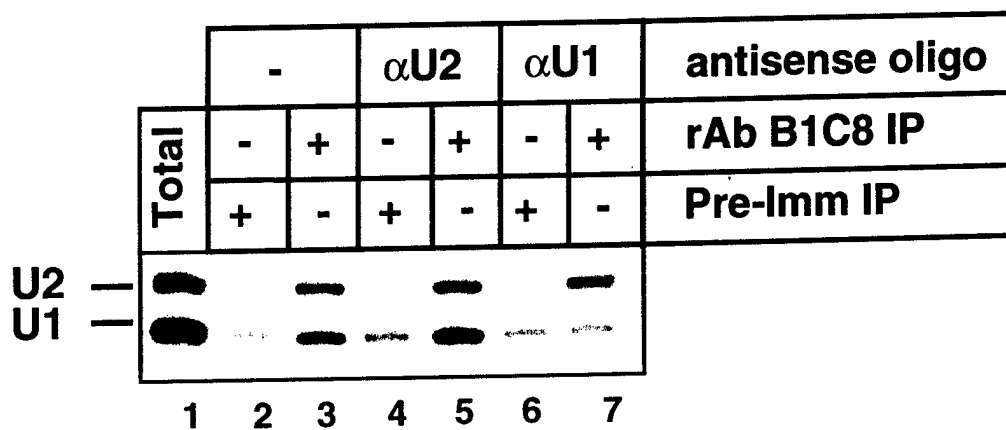|		-		αU2		αU1		antisense oligo
Total		-	+	-	+	-	+	**rAb B1C8 IP**
		+	-	+	-	+	-	**Pre-Imm IP**

U2 —
U1 —

1 2 3 4 5 6 7
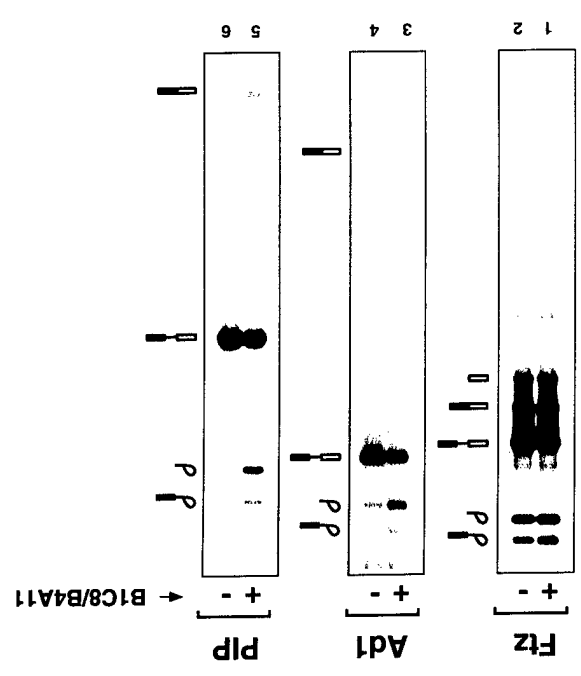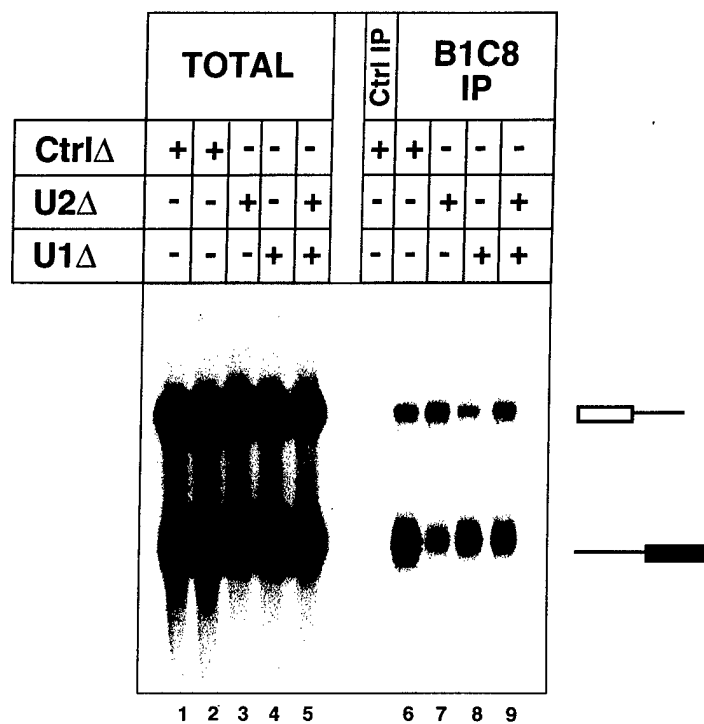
Figure Figure 6C (2)

Figure 7

Figure 8